DOE/NSF U.S. LHC PROGRAM OFFICE
APRIL 6, 2006

REPORT ON THE JOINT DOE/NSF REVIEW OF U.S. LHC SOFTWARE AND COMPUTING

AT THE UNIVERSITY OF CALIFORNIA,
SAN DIEGO
7-9 FEBRUARY 2006

REPORT TO THE U.S. DEPARTMENT OF ENERGY AND
THE NATIONAL SCIENCE FOUNDATION

TABLE OF CONTENTS

EXECUTIVE SUMMARY

The external reviewers of the U.S. CMS and U.S. ATLAS Software and Computing (S&C) programs (see Appendices A, B and C) commended the U.S. teams for their excellent progress since the previous comprehensive review of February 2005, and for the clear and concise way that the management teams addressed former concerns, and presented the status of their respective programs.

The reviewers offered several observations and recommendations common to the two experiments. They noted that both U.S. collaborations have sound S&C management structures and that they are well represented in leadership positions in their respective international collaborations. They also recommended that US ATLAS and U.S. CMS continue assessing their scientific goals and the implications of these for computing resources. Priorities should be re-visited regularly to assure that the most urgent computing needs can be accommodated within agency funding guidelines. The committee also recommended that both collaborations work together to clarify the reasons for the differences in their respective computing requirements in order to minimize any impact on overall U.S. LHC goals.

The reviewers also felt that incremental costs of computing resulting from increased U.S. participation in the LHC experiments should not be allowed to foreclose the opportunity for attracting new U.S. collaborators, but that U.S. ATLAS and U.S. CMS should rather consider this as a challenge to develop a reliable model for absorbing incremental costs, and seek sources of additional funding, as required.

At this stage of the S&C program, the focus is on deploying the required technologies, in particular, the facilities, networking and grid infrastructure that will facilitate analysis at remote sites. The projections for required human resources at both collaborations were deemed appropriate for the S&C effort at hand. In fact, the current ramp-up in personnel at the Tier-1 and Tier-2 facilities and in the support of analysis is aimed to provide successful LHC operations. Both collaborations, and in particular CMS, are expecting a large influx of US physicists who are redirecting their efforts from other collider experiments, and it is expected that many of these new collaborators, especially the university scientists, will contribute to the S&C efforts.

A serious concern expressed by the reviewers was that, given the scope defined by the U.S. teams, the DOE/NSF funding guidelines do not provide sufficient flexibility for the collaborations to respond quickly to unanticipated problems. Because expertise is not an "off-the-shelf" commodity, the collaborations have chosen to place a higher priority on maintaining the necessary workforce than on equipment purchases. In the case of CMS, the reviewers feared that the shortfall in international pledges for Tier-1 facilities may also reduce capabilities for analysis by U.S. groups. While this shortfall in pledges is a major concern, the reviewers felt that it has to be addressed and resolved by the international CMS collaboration.

Both experiments have dependencies on external grid middleware, which is being maintained and provided by the currently-unfunded Open Science Grid (OSG). If long-term funding for OSG cannot be assured, both collaborations would have to extend their scope and assume responsibility for the orphaned middleware components, placing additional pressures on their budgets.

The CMS LHC Physics Center (LPC) at Fermilab has been a success due to the strong commitment by U.S. CMS management and due to the enthusiastic adoption of this model by the U.S. collaboration. One example of LPC success was the re-design and implementation of the CMS core framework and the Event/Data Model. Although there was concern about the difficulty of this project at our two previous reviews, U.S. CMS took ownership of this problem, assembled an experienced development team, and delivered a successful product. Nevertheless, much work still remains to be done in integrating the new packages with the existing reconstruction software.

U.S. ATLAS has recently developed an analysis-support model that relies on distributed centers of experts across the U.S., with a core group of experienced individuals providing the required expertise. While this model has been endorsed by the U.S. ATLAS collaboration, the committee was concerned that the implementation of this plan relied on funding that has yet to be identified.

The collaborations presented scenarios for a 10% reduction in current funding guidelines that included descoping of their computing facilities, possibly defaulting on their international commitments as well as the option of reducing support for domestic analysis. The committee felt that the S&C budgets were already at a minimal level, and that reductions would not be acceptable. For example, in the case of ATLAS, assuming that no calls for management reserve are granted, current S&C plans foresee a $3M shortfall in 2008. Any additional shortfall would jeopardize the important role U.S. ATLAS plays within the collaboration, and would endanger the opportunities for direct participation in scientific discovery by the U.S. team.

1 INTRODUCTION

Software and Computing will serve as a key element for enabling the science of U.S. physicists at the LHC. The unprecedented data rates and the distributed nature of collaborative analyses pose logistic challenges of access to data that can only be tackled with an effective and broad software and computing infrastructure. A viable system is particularly important for U.S. physicists, who must overcome the disadvantages of geographic separation from their experiments.

The U.S. Department of Energy and the National Science Foundation, partner agencies supporting U.S. participation in the LHC, have recognized the importance of Software and Computing (S&C) to the success of the U.S. investment in the LHC. The U.S. collaborations have established S&C programs within the context of their Research Program Management Plans, and in order to monitor progress and establish support guidelines, the agencies regularly review the work of the collaborations.

At the request of the DOE/NSF Joint Oversight Group (JOG), a review of the ATLAS and CMS S&C programs was held on February 7-9, 2006 at the University of California, San Diego (UCSD). The review covered the general areas of Management and External Interactions, Facilities and Grids, and Core Software. A charge to the review committee (Appendix A) was prepared by the JOG, and distributed to the panel members and the collaborations prior to the meeting at UCSD. The members of the review panel were chosen for their expertise in software and computing, and came from U.S. universities, U.S. laboratories, and European institutes, outside the LHC community. The list of participating panel members is included in Appendix B. The agenda of the review is included in Appendix C. The presentations by the collaborations, the full agenda, and additional material can be found at http://www.disun.org/2006LHCReview.html.

The expert reviewers and the collaborations were instructed to address the progress in S&C to-date, and to assess the needs and plans for the period FY2007–FY2010. This report, prepared by the external reviewers and edited by the agencies, summarizes the findings, observations, and recommendations of the panel members. The observations and recommendations are based on the presentations, on background material provided prior to the review, on detailed discussions during the parallel sessions, and on answers to questions posed by the panel. The panel members also deliberated on their observations and recommendations in executive sessions in the presence of representatives from the DOE and NSF, and then presented their preliminary findings in close-out sessions.to the collaborations.

UCSD, as host for this meeting, provided invaluable logistical support, which contributed greatly to the success of the review

2 MANAGEMENT

2.1 U.S. CMS

The U.S. CMS Software and Computing (S&C) management effort was summarized in a presentation by L. Bauerdick, and in breakout meetings with L. Bauerdick, I. Fisk and D. Green.

2.1.1 KEY FINDINGS

The U.S. CMS management team for Software and Computing was put in place in 2005, and is reasonably well balanced between FNAL and university members. The team has met most of this past years's deliverables and has kept within its scope .

There have been notable success in meeting high-level milestones, including participation in the recent test of the central magnet and through a demonstration that the production system can accommodate increased complexity. The goals of Service Challenge 3 (SC3) included testing a quasi-production system that supplied the basic functionality for the major foreseen patterns for data transfer and data serving. Much of the focus of SC3 involved debugging components that had high rates of failure, dealing with shortcomings of components known to be inadequate (such as the old framework) and testing those elements of the grid that led to a fragile operating system that was labor intensive. Success was achieved in the reasonable throughputs obtained in T0 to T1 transfer rates, and in demonstrating excellent performance of storage systems at several sites, including Tier-2 centers. The US Tier-1 site, considered a model facility within CMS, hosts the heavily accessed JETMET Monte Carlo events.

The Analysis Environment specified in the Physics TDR was released on schedule in July 2005, thereby providing access of al U.S. CMS to basic analysis steps and submission of jobs for analysis at remote centers. The release of the Data Management Prototype, targeted for July 2005, has been delayed, although a test version is available at selected sites. Much of the delay was caused by the loss of key developers, and more effort will be added to this task. Hiring staff at FNAL and the Tier-2 sites is progressing, but at times at a slower pace than desirable.

The development of the latest CMS framework was proposed in 2005 to address some severe deficiencies found in the previous framework. The basic code is now completed and tested. Expert consultants are being provided to assist CMS developers in converting their software to use with the new framework, to which and several additional features are still to be added.

U.S. CMS contributed significantly to the CMS Computing Technical Design Report, which provides a thorough assessment of needed resources, as determined through an evaluation of the computing model. The Report includes a list of services needed at the

baseline level, and describes the process for increasing functionality through incremental improvements. The main computing milestones involve primarily integration tasks. To optimize the use of resources, data is placed at sites either through explicit CMS policy or via skim processes. Data sets are organized around trigger selections, with "raw" and reconstructed outputs distributed as custodial copies to Tier-1 centers, and full analysis copies to each Tier-1, with partial copies available at Tier-2s.

There is a 40% projected shortfall in global Tier-1 resource needs from absence of pledges in the international collaboration. This problem is likely to persist well past the start of the LHC data taking. The origin of the gap is that CMS has member nations that cannot host a Tier-1 faculty, but the experiment has no provision for those nations to contribute to other CMS Tier-1 sites.

2.1.2 OBSERVATIONS

U.S. CMS is strongly committed to the success of international CMS. The U.S. S&C management team is working very effectively with its partners, with an appropriate number of U.S. members serving in the management of international CMS Computing. The overlap in responsibilities between international and U.S. management roles occasionally results in unclear distinctions about the separation of goals and contributions of U.S. CMS S&C and those of the more global effort. This can lead to an overly protective attitude that can miss critical issues. The process of refining the goals for the cosmic-ray studies and the test of the central magnet should provide a good opportunity to bring some clarity to the dual management roles.

The U.S.-led success in developing the CMS framework was achieved by making productive use of the expertise within the FNAL Computing Division. The new framework has facilitated contributions to the software that can be made by physicists, and is a fine example of how U.S. CMS, working within international CMS, can solve critical problems in S&C.

The LHCC has noted that there is a factor of two difference in computing requirements between the global CMS and ATLAS efforts, yet, there does not appear to be any fundamental difference in needs of the two analyses. Following review of the Computing TDRs, both collaborations have been asked by the LHCC to come to an understanding for the source of this difference in the respective computing requirements.

U.S. CMS is very concerned about the funding of CMS computing resources because, despite the fact that the needs for CMS are projected to be much lower than those of ATLAS, CMS has a global 40% shortfall in its Tier-1 resources. The CMS collaboration will have to look to internal processes for prioritization and allocation of resources to cover part of this shortfall. This will be essential for helping U.S. CMS build a strong case for any impact of the remaining shortfall on its ability to succeed in scientific goals.

The U.S. CMS contribution to Tier-1 currently stands at 28% of needed resources. However, relative to internationally pledged resources, the US will contribute 36% of the CPU, 39% of the disk and 58% of the tape resources. For Monte Carlo generation at Tier-2 sites, the U.S. has pledged 28% of its resources to international CMS. There are 7 Tier-2 sites, of which 4 are associated with DISUN. In addition, the LPC-CAF at FNAL represents a significant analysis resource, comparable to ~5.5 Tier-2 sites. Co-located with the Tier-1 facility, the LPC-CAF would provide the computing to support 220 physicists while 7 US Tier-2 sites would support 280 physicists. If confronted with a 10% reduction in funding, U.S. CMS would likely choose to reduce the number of Tier-2 centers, with a resulting loss of leveraging. U.S. CMS expects eventual funding of the OSG by DOE and NSF for providing advanced grid functionality, and funding from DISUN for some of the Tier-2 support.

The number of U.S. collaborators in CMS increased by 11% in 2005, and is anticipated to continue growing. This implies an increase in U.S. CMS computing needs (and costs) to enable effective participation in CMS for the additional U.S. collaborators. U.S. CMS maintains that the policy for membership in the CMS collaboration is an international CMS responsibility, and chose not present the conditions under which new U.S. groups should be accepted into CMS. The observed substantial growth in participation makes it crucial to set priorities for the use of computing resources.

U.S. CMS proposed that pledges to the global effort for U.S. Tier-1 computing should be increased as the size of the U.S. participation continues to grow. The review committee was not convinced, however, that an increase in pledges in proportion to the growth of the U.S. team was warranted, since the original size of the Tier-1 was driven by global CMS Tier-1 computing needs, which are not expected to be very sensitive to the relatively small increases in U.S. membership.

It was noted that U.S. university groups are making significant contributions to the software effort, with new groups being well integrated into CMS, and encouraged to contribute to the S&C work plan.

CMS management expressed some concern about the timely availability of robust grid middleware that is needed to support production at LHC turn-on. The CMS managers have pursued a prudent policy of ensuring that their production pipelines are able to interface to current grid middleware or, in case CMS encounters difficulties, to custom CMS-specific tools developed for the collaboration to exploit its Tier-1 and Tier-2 resources.

2.1.3 RECOMMENDATIONS

The Committee has the following recommendations for U.S. CMS management:

- We recommend that U.S. CMS reexamine allocation of computing resources between Tier-1 and Tier-2 centers (and LPC/CAF) so as to absorb part of the expected shortfall in resources at the Tier-1 center.
- We recommend that the collaboration reexamine and adjust its scientific priorities to accommodate any remaining shortfall in Tier-1 resources.
- We recommend that the U.S. CMS collaboration remain open to new collaborators, in spite of limited Tier-1 and Tier-2 resources. U.S. CMS should continue to assess its priorities and computing needs, pursue other funding opportunities, and work with the funding agencies to obtain adequate resources for its Tier-1 and Tier-2 facilities.
- We recommend that U.S. CMS S&C should continue to hold to current scope.
- CMS should continue to refine its computing model by optimizing performance and resource requirements.
- We recommend that U.S. CMS continue its prudent and conservative approach to reliance on grid middleware.
- We recommend U.S. CMS and U.S. ATLAS work together on reconciling the estimated computing costs per additional collaborator.
- U.S. CMS should work with U.S. ATLAS to clarify the reasons for the differences in their respective computing requirements in order to minimize any impact on overall U.S. LHC goals, and that they address the questions posed to the collaborations by the LHCC.

2.2 U.S. ATLAS

The U.S. ATLAS management team for Software and Computing of J.Shank et al presented the status of their effort during the plenary and parallel sessions.

2.2.1 KEY FINDINGS

U.S. ATLAS S&C has achieved a very important milestone with the completion of the Computing Technical Design Report in June 2005. This document contains a thorough evaluation of the experiment's computing needs. It was reviewed by the LHCC, and memoranda of understanding (MOU) are now ready to be signed.

Activities related to the Rome physics workshop (June 2005) and the 2nd North American Physics Workshop in Toronto (Aug 2005) were discussed at the presentations. Much of the total ATLAS GRID production for Monte Carlo events for these meetings was provided through U.S. ATLAS facilities. Twenty different sites worldwide participated in this effort, with the three Tier-2 sites and the Tier-1 site responsible for more 75% of the events. During the production run, the "Southwest Tier-2", a consortium of three universities, surpassed in performance even the BNL Tier-1. Despite the success, several weaknesses were identified during this effort: The ATLAS GRID production system exhibited serious problems for increased throughput. In addition, it proved to be very labor intensive, requiring continuous attention.. Contention for resources required manual

intervention, and it became quite apparent that there was an urgent need for an organized approach to support user analysis. The management reacted to these problems through several means, as indicated below.

The "PanDA" project was started with the goal of developing a new distributed analysis and production system, replacing the previous "CAPONE" system. The development of an analysis-support model was initiated and agreed upon within the U.S. ATLAS collaboration. The implementation of this model is now under way. The organizational structure in the "Software and Analysis Support" activity was adjusted accordingly. This meant the creation of two new activities at Level 3, namely "Distributed Software", headed by Torre Wenaus and "Analysis Support", headed by Stéphane Wilocq. The PanDA development takes place within Distributed Software, and the implementation of the support for user analysis is executed by Analysis Support.

A resource-allocation committee (RAC) was set up in order to deal with contention for resources in a rational and transparent way. The RAC brings together top management, representatives from the physics groups and the key players on the provider side.

In response to the recommendations of last year's review panel, a change-control procedure was introduced in the ATLAS Software and Computing project in 2005. The PanDA effort was, in fact, added via the new change-control process.

Activities during 2006 will be driven by a three-phased staged Computing System Commissioning plan (CSC). The CSC is, in effect, a redefinition of the previous Data Challenge 3 activity. Phase 1 involves the commissioning of the general infrastructure and individual subsystems. This started in 2004, and will continue through the spring of 2007. During Phase 2, which runs in parallel with Phase 1, subsystems will be combined and commissioned to evolve into an integrated experiment. Finally, Phase 3 will start in the spring of 2007, and will morph from running with cosmic rays and single beam to first collisions.

2.2.2 OBSERVATIONS

The committee commends the U.S. ATLAS Software and Computing team on the excellent progress achieved during the past year, and, in particular, for the thorough evaluation of ATLAS computing needs and the timely delivery of the ATLAS computing TDR. This has been one of the most important steps on the way to successful data taking in 2007. Nevertheless, ATLAS should keep improving its computing model and continue efforts to optimize its performance and resources.

The committee was impressed by the success of the multi-site Tier-2 model, as demonstrated during the grid production leading up to the Rome workshop. This model was a concern during previous reviews, but recent experience has shown that it is indeed

workable. U.S. ATLAS Tier-2 centers should be commended for providing substantial leveraged support from the participating universities.

The committee also notes that U.S. ATLAS S&C management has been very responsive to the comments of the last year's review panel. As was suggested, U.S. ATLAS has introduced change-control, and has agreed on a common model in support of analysis. These developments represent important steps forward for U.S. ATLAS. However, the implementation of the model for the support of analysis requires funds in 2008 and beyond that have yet to be identified. The committee feels that this is a major concern, because support of user analysis will be essential to the success of the physics program. This was already noted two years ago during the 2004 review at Fermilab. Currently, it appears unlikely that the missing funds will materialize. The implementation plan for the model therefore needs further optimization with the goal to accommodate it within the present U.S. ATLAS S&C budget. The committee also notes that the contention presented for the remaining management reserve in U.S. ATLAS is an internal matter that should be resolved accordingly.

The committee supports the planned transition from construction to operations during first beams in 2007 in the form of a three-phase commissioning of the computing system. This approach is reasonable and appropriate. However, the committee wishes to emphasize the paramount importance of the cosmic runs in this planning. As it was presented, it appears that complete agreement on the scope and timing of these runs has not been achieved within International ATLAS. This is a concern, since the cosmic runs present the first opportunity to test the full data chain from trigger, via acquisition, archiving, reconstruction and distribution to analysis, at low rates, and therefore presents ideal conditions for preliminary debugging.

The committee commends U.S. ATLAS S&C on the formation of the resource allocation committee (RAC). This body can be of major importance during the commissioning of the computing system, and during the difficult conditions of first beams, when quick turnaround is required and demands will likely outpace available resources. The committee believes that the RAC will have to act flexibly and quickly, and its present large size and composition may therefore not be optimal.

The Computing System Commissioning plan is sufficiently advanced so that a cosmic-ray run in May 2006 could provide an extremely valuable first test of the entire system. Current planning will require the entire management reserve of U.S. ATLAS in 2008 and beyond to be applied to S&C. The projected budget shortfall in U.S. ATLAS S&C will be of the order of $3M per year. The committee believes the planned S&C budget, as presented, is already sub-critical and at a minimal sustainable level. Not fulfilling the call for the entire reserve would jeopardize the important role U.S. ATLAS plays within the collaboration and would endanger the opportunities for physics discovery by the U.S. team.

Finally, there was discussion of the incremental computing costs associated with the likely growth of the U.S. ATLAS collaboration. These costs are in addition to "Category A" that U.S. ATLAS would have to pay per collaborator. The computing costs originate from the need for additional capacity at both the Tier-1 and Tier-2 centers in order to support a larger number of collaborators. Current estimates are that each collaborator would cost ~11k$/year. However, since additional funds are unlikely to materialize, it is not clear how this cost can be covered. U.S. ATLAS, while keeping the doors to new U.S. collaborators open, goes through an evaluation process before recommending the admission of additional US collaborators to ATLAS. The committee finds the above cost estimates reasonable, and agrees with U.S. ATLAS that while the collaboration needs to remain open, a strategy for admitting new collaborators is also required. In view of the fact that U.S. ATLAS is likely to grow, the committee encourages the collaboration to continue to assess their physics priorities and the corresponding computing needs, to pursue other funding opportunities and to work with the funding agencies to obtain adequate resources.

2.2.3 RECOMMENDATIONS

The committee made the following recommendations:

- ATLAS should continue to refine its computing model, and optimize its performance and resources.
- The implementation of the plan for support of user analysis must be adapted to the present U.S. ATLAS S&C budget guidelines.
- The contention over the U.S. ATLAS management reserve is an internal matter that should be resolved accordingly.
- U.S. ATLAS should encourage international ATLAS to define a plan for the complete cosmic-ray run that would exercise the trigger, data acquisition, and software reconstruction chain.
- The size and composition of the Resource Allocation Committee should be optimized as experience accrues.
- The U.S. ATLAS collaboration should remain open to new collaborators; however, ATLAS should continue to develop a strategy that does not compromise the success of the project due to insufficient computing resources.
- Since the collaboration is likely to grow, U.S. ATLAS should continuously assess the physics priorities and the corresponding computing needs, to pursue other funding opportunities and to work with the funding agencies to obtain adequate resources.
- U.S. ATLAS should work with U.S. CMS to clarify the reasons for the differences in their respective computing requirements in order to minimize any impact in overall US LHC physics. Management should also address the questions posed to the collaborations by the LHCC.

3 FACILITIES AND GRIDS

3.1 U.S. CMS

The U.S. CMS Facilities and Grids were discussed in plenary talks by Bauerdick (Overview), K. Bloom (Tier-2), and F. Wuerthwein (DISUN). In the breakout session, there were additional presentations by J. Bakken (Tier-1), D. Petravick (Networking), and R. Pordes (OSG).

3.1.1 KEY FINDINGS

The capacity of the Tier-1 center continues to ramp up. It already has 25% of the planned CPU capacity and 100 Terabytes of storage installed. A significant fraction of the Monte Carlo events generated for the CMS TDR can be accessed at FNAL. In 2006, the CPU capacity of the Tier-1 center will be doubled to ~50% of total capacity, and most of the remaining CPU will be purchased in 2007.

To understand issues of scaling with event rate, the storage capacity of the Tier-1 center will be increased to 1.5 Petabytes in 2006. Network connectivity at the Tier-1 facility continues to improve and is well on the way to having an aggregate of 60 Gbps connectivity, including redundant links.

The LPC-CAF is ramping up and will have 0.5 Petabytes of storage in 2007. The LPC is beginning to support users, and the CAF(???) now has ~450 registered users. The CMS service challenges used the available production hardware, which was installed without affecting user access to the facility. The target for data transfer of the service challenge was met, despite the presence of several identified bottlenecks in rates for data transfer. I/O queues were implemented on the fly (???) to support sustained transfer rates of sufficiently high bandwidth. Support personnel recently lost to industry have made it difficult to keep the facility staffed with experienced people.

Countries representing only 60% of the author list have contributed to Tier-1 facilities of global CMS, with their contributing share defined by the fraction of those pledging support. This has produced a deficit of 40% in pledges to Tier-1 resources caused by the missing 40% of CMS, and presents an important issue that international CMS must resolve rapidly, or suffer serious constraints in computing.

Four new Tier-2 centers were selected in 2005 (MIT, Nebraska, Purdue, and Wisconsin) and the successful DISUN proposal to NSF funded an additional three centers (UCSD, Caltech, and Florida). As of December 2005, these Tier-2 sites had reached 25% of the CPU and 10% of the storage planned for the Tier-2s. Networking bandwidth to the Tier-2 sites currently ranges from 155Mbps to 10Gbps, and there are plans for upgrading most of these sites to shared 10 Gbps transfer rates. All Tier-2s have pledged to have at least 2.5 Gbps capacity by 2008.

For the Tier-2 centers that participated in SC3, ~50% of the jobs that ran completed successfully. Most of the failures were due to flaws in the core software that should be addressed by the recent deployment of the new framework and of a more robust bookkeeping system. However, 16% of the failures were due to grid middleware, which is presently affected by serious limitations in scaling with rate.

The grid middleware is now integrated into operations, and is maintained and provided by the Open Science Grid's (OSG) "volunteer effort" in the form of the Virtual Data Toolkit. Many of the issues with interoperability with the LCG software have been resolved and the LCG and OSG groups are cooperating to address remaining issues.

U.S. CMS is closely coupled to the management of the Open Science Grid, and has key personnel in leadership roles of the OSG. We note that in the past year OSG has made the transition from the demonstration of Grid3 to a national-scale production cyber infrastructure with a first software release (OSG 0.2.1) in July 2005 and the second major release (OSG 0.4.0) in January 2006. U.S. CMS relies on several key services that are currently provided by OSG.

3.1.2 OBSERVATIONS

The U.S. CMS Tier-1 facility will be doubling its CPU capacity over the next year. The storage capacity at the Tier-1 will also increase dramatically in 2006. This ramp up will be critical to gaining an understanding of the scaling of response with the size of the storage system. The support from Fermilab Computing Division has been essential to the successful installation of the Tier-1 facility. The LPC is beginning to support users from afar, and it now has about 450 registered user names. However, only a small fraction of these users are submitting most of the jobs.

The shortfall of global CMS pledges to Tier-1 facilities may reduce analysis capabilities to a scale commensurate with available capacity. U.S. CMS has not been able to affect the global Tier-1 shortfall. A 10% reduction in the U.S. funding would exacerbate this situation since it would delay installation of resources at the U.S. Tier-1 facility (and not even meet the U.S. pledge). It would also delay testing of the production system at full scale, which would leave little time to debug any new problems introduced by running the eventual system. A 10% reduction in funds would not affect personnel, since U.S. CMS considers the personnel level to already be at bare minimum.

Progress in identifying and addressing failure rates of the middleware has been slow so far, however, significant improvement is expected through deployment of the new strategy for managing.

The fragility and scalability of the grid middleware-based job submission system is a concern and should be addressed. Although workarounds such as gridmonitor are in place, the underlying job submission components such as GRAM have to be fixed to provide the needed scaling. U.S. CMS is embracing enthusiastically cooperation with U.S. ATLAS and other science projects to develop and use common middleware through the Open Science Grid. The OSG is currently providing to U.S. CMS the VDT software, software validation, integration, and operations. It is apparent that U.S. CMS is making significant contributions to, and is receiving significant benefit from, its participation and reliance upon OSG. Many of the functions currently provided to U.S. CMS by OSG would have to be addressed by U.S. CMS if OSG is not funded. In addition, U.S. CMS would have to increase staff if it had to provide these functions.

Good progress on selection, funding and commissioning of the Tier-2 sites has been made resulting in 7 Tier-2s, which have reached ~25% complexity and plan to reach 50% complexity by the end of 2006. The job failure rate at the Tier-2 centers was unevenly distributed, and some sites had an alarming error rate. Much of this is likely an artifact of the start-up of these centers, but should be aggressively pursued and reduced.

Overall, the planned increase of computing resources and network throughput, as well as the projected level of operational support, seem well matched to the analysis needs of the current U.S. CMS community, as outlined in the current CMS computing model. However, no contingency remains for coping with unexpected shortages and/or the large expansion of the U.S. CMS group.

3.1.3 RECOMMENDATIONS

The committee makes the following recommendations:

- To gain experience with a large-scale system, U.S. CMS should not delay the current plan to purchase equipment for its Tier-1 center.
- U.S. CMS should participate in CMS discussions to address the Tier-1 shortfall in resources, explore re-optimization of resources between Tier-1 and Tier-2 centers, and prioritize its scientific goals should constraints be forced by the shortfall.
- U.S. CMS should request the grid middleware projects to hire personnel for this year's service challenges to address the problems of scaling with data rate.
- U.S. CMS should devise a metric that summarizes the overall progress in reaching the milestones of the Tier-1 and Tier-2 preparatory work that are similar to the monitoring plots that CMS produces for detector construction.

3.2 U.S. ATLAS

U.S. ATLAS Facilities and Grids were discussed in presentations by B. Gibbard (Tier1 – plenary and parallel), R. Gardner (Tier2 – parallel, OSG/WLCG on behalf of J. Huth –

parallel), S. McKee (Networking – parallel) as well as by J. Shank in the overview of the S&C activity.

3.2.1 KEY FINDINGS

The projected hardware capacity for the ATLAS Tier-1 facility in 2008 is sufficient to meet the resources pledged to the WLCG. An additional 50% CPU, 50% Disk, and 100% Tape capacity has been reserved for exclusive use of U.S. scientists. The plans for 2008 include adequate staffing at the Tier-1 center. However, this projection includes an assumption of receiving $2M from management reserve (MR) in 2008. If a 10% reduction in base funding occurs in 2008, this projection would require $3M from MR.

There were two "reduced funding" scenarios presented. In the scenario assuming no MR, the capacity profile for the Tier-1 just meets that pledged to the WLCG, without providing additional capacity to US scientists. While capacity for tape is retained, there is some reduction in bandwidth to tape. In this scenario, the level of personnel is frozen at 16 FTE (the 2007 level) rather than increasing to 20 FTE in FY'07. Assuming no MR and a 10% reduction in base funding, the capacity falls below the level pledged to WLCG after 2008. For the latter situation, there is also a reduction in personnel to 14 FTE for 2008 and beyond.

The Tier-1 facility has been successfully operating 200 TB of dCache-managed disk space, distributed across 400 computer nodes for a full year.

The Tier-1 network connection to the NYC MANLAN, and consequently to LHCNet and CERN, was upgraded to 10 Gbps, and was scheduled to be upgraded to 2 x 10 Gbps ("2 λ") in March 2006. It was noted that this is not a redundant connection, and can be lost for hours/days by a single incident. The 2 λ connection meets the ATLAS bandwidth requirements for the start of LHC data taking.

U.S. ATLAS has established three Tier-2 centers (Northeast, Midwest, and Southwest) and is planning to establish two additional centers in 2006/2007, with selection planned by July 2006. The Northeast Tier-2 is located at Boston University. The Midwest Tier-2 is distributed between the University of Chicago and Indiana University. The Southwest Tier-2 is distributed between Oklahoma U., U. Texas-Arlington, and U. New Mexico. The projected aggregate capacity of the five Tier-2 centers in 2008 meets the amount pledged to the WLCG, with an additional 50% (CPU) and 100% (Disk) of pledged capacity reserved for US scientists. It was mentioned that, in a "reduced funding" scenario, U.S. ATLAS would consider establishing one rather than two new centers, choosing a total of four centers by 2008. The committee feels that eliminating a Tier-2 center would not be cost effective since ATLAS benefits from extensive leveraging at the universities. However, this has not yet been fully analyzed by U.S. ATLAS.

The production of simulated data was carried out for Data Challenge 2 and the Rome Physics workshop using Grid3/OSG in the U.S. and other international grids. This was considered a great success, but experience showed it to be too labor intensive. This led to the development of the PanDA system (covered in the following section of this report). Preliminary results of the effectiveness of using the PanDA system for managing grid-based production are encouraging, both in terms of failure rate and personnel requirements.

A planned transition from building up of the facility to operating the Tier-1 at BNL shows a steady-state staffing level of 20 FTE to be reached in 2007.

A working group was established by the U. Chicago Tier-2 staff and the Argonne Analysis Support Center to consider issues of user support of analysis at Tier-2 centers.

The U.S. ATLAS effort at the Tier-1 center is involved in the cyber security activities related to user management and supporting role-based authorization for grid users. They have developed and support the GUMS (Grid User Management System) for OSG and its interactions with other "Privilege" project components. It was noted that network and host-level cyber security for the BNL facilities are provided by the BNL IT division.

Condor and Globus developers have received significant funding from NSF to support their middleware (Condor - $4.2M over 3 years; Globus - $13.3M over 5 years), and U.S. ATLAS and U.S. CMS personnel, among others, are collaborating on a proposal to the NSF and DOE for the support of the Open Science Grid.

3.2.2 OBSERVATIONS

Brookhaven National Lab has started charging U.S. ATLAS for power costs incurred by the Tier-1 facility. This expense was not originally anticipated in the facility cost projections. In light of the uncertainty in the future cost of power, the projections for power consumption presented at the review may be difficult to gauge, but U.S. ATLAS is investigating this matter further.

The plan for the ramp-up in Tier-1 personnel seemed reasonable to the review panel. However, a reduction in personnel from an inability to fulfill a $2M request for Management Reserve in FY08 and beyond would introduce serious risk in the ability of the Tier-1 facility to provide the required functionality.

The U.S. ATLAS Tier-1 has chosen to deploy a disk-based event storage using dCache-managed disks that are distributed on many compute nodes, worldwide. This provides reduced hardware costs compared to traditional disk arrays, but may bring additional risk from the application at unprecedented high data rates.

Three Tier-2 centers have been selected and established by U.S. ATLAS. Each of these involves a multi-site installation. Some of these Tier-2 centers have achieved significant leveraging of resources. Discussion is in progress to address physics analysis and user support, as well as the relationship between a Tier-2 center and an Analysis Support Center.

The BNL Tier-1's WAN connection is currently not redundant. Although there are ongoing efforts to ameliorate this deficiency, there is as yet no concrete plan to add a redundant network connection. The currently planned U.S. ATLAS connectivity for Tier-2 centers is more than adequate, and is expected to achieve 10 Gbps rates at each site by the end of 2007.

U.S. ATLAS pointed out that the failure of PanDA-controlled job submission is significantly reduced compared to the system it replaced, and it has provided the ability to handle significantly higher rates of job submission.

The OSG "volunteer effort" is currently providing VDT software, software validation, integration, and operations, upon which U.S. ATLAS depends. The committee noted that many of these functions would have to be assumed by U.S. ATLAS if the Open Science Grid is not funded.

3.2.3 RECOMMENDATIONS

The committee recommended that:

- Further efforts be undertaken to define the relationship and responsibilities of physicists at the Tier-2 centers and those at the the Analysis Support Centers.
- U.S. ATLAS should try to assure that personnel at grid middleware projects become engaged in this year's service challenges, and especially in issues pertaining to scaling.
- U.S. ATLAS should pursue acquisition of redundant network connectivity to the BNL Tier-1 center.
- U.S. ATLAS should test the scaling properties of dCache-managed distributed disk systems.

3.3 ADDITIONAL OBSERVATIONS ON ATLAS AND CMS

The committee observed that the cost and capacity profiles of U.S. ATLAS and U.S. CMS Tier-1 facilities are notably different. This appears to be due mainly to the different strategies for achieving the required disk capacity and performance goals. U.S. CMS has chosen a conservative approach in deploying disk space by using dedicated file servers while U.S. ATLAS has chosen to use a disk system distributed across the computing nodes. Another notable difference is that U.S. CMS is deploying significantly more disk

and CPU during 2006 and 2007 while U.S. ATLAS plans on deploying more disk and CPU in 2008. The overall effect of these choices is that U.S. ATLAS achieves greater capacity at lower cost but is assuming somewhat greater risk compared to U.S. CMS.

4 CORE SOFTWARE

4.1 U.S. CMS

4.1.1 KEY FINDINGS

U.S. CMS has taken on significant management responsibilities for core software within global CMS. It has also taken responsibility for implementing, deploying and supporting a new software application framework, and a new Event Data Model (EDM). The need for replacing these key elements was identified in the 2004 Data Challenge, and the participation of U.S. CMS in providing the replacement was discussed at previous S&C reviews.

The LHC Physics Center (LPC) at Fermilab is continuing to develop into a bone fide U.S. CMS analysis center, as well as a meeting place for U.S. CMS physicists.

U.S. CMS has begun developing new reconstruction algorithms (particularly tracking) that address shortcomings of the existing software, and which are relevant to the upcoming Computing, Software, and Analysis challenge ("CSA 2006") software tests scheduled for this year.

4.1.2 OBSERVATIONS

As indicated previously, there is extensive participation by members of U.S. CMS in the management of the global CMS software. Given their extensive experience in software development and management for the Tevatron, this is commendable and appropriate.

The re-engineering of the framework and development of the Event Data Model (EDM) has progressed rapidly since the last review. The framework and EDM designs borrow heavily from the experience at the Tevatron, and particularly from CDF, which gives them a solid foundation and leverages the experience of the developers. The implementation and deployment of these components is on schedule, and their acceptance by the experiment is rapidly solidifying. The benefits to the collaboration are already starting to manifest themselves in the tools of the new framework (e.g., "leak" checkers) leveraged from their origins at CDF. Members of international CMS are also reporting success in using the new framework. In many respects, the replacement of the framework and the EDM can be considered as completed. In particular, the committee was presented with detailed descriptions of how U.S. CMS core software efforts have migrated from development to support and maintenance. Nevertheless, the associated migration and integration of all CMS software modules into the new framework is not yet finished. While this migration is not a direct responsibility of U.S. CMS, it may affect the ultimate success of its adoption, and the effectiveness of the upcoming data challenges. The committee felt it important to validate the performance of the new framework and EDM in the upcoming CSA 2006 end-to-end tests.

The LPC at Fermilab is attracting more users and is developing a more mature management structure. The LPC appears to have succeeded in its primary goal of providing a central location for support of U.S. CMS software. Current developments in tracking and other reconstruction algorithms show that there is an appropriate focus on activities likely to have major impact in the analysis of initial data.

4.1.3 RECOMMENDATIONS

The committee recommended:

- U.S. CMS should develop a set of specific measures to be monitored during the CSA 2006 tests to confirm the effectiveness and suitability of the new framework and EDM.

- U.S. CMS should define the general goals of the upcoming end-to-end test (CSA 2006) and the associated U.S. CMS responsibilities.

- Now that the focus has shifted from design and implementation of the new framework and EDM, U.S. CMS should prepare a more detailed plan for activities of their software developers. In particular, the U.S. collaboration must choose between extending development in the framework towards additional functionality and new efforts such as development of reconstruction and analysis algorithms.

- While the existing staffing plan for maintenance and support activities seems appropriate, the review panel recommended that U.S. CMS prepare a detailed plan for the upcoming end-to-end test that would buffer key developers from the expected onslaught of questions from the user community.

4.2 U.S. ATLAS

4.2.1 KEY FINDINGS

A large-scale exercise, Data Challenge 2 (DC2), was carried out in early 2004, and deemed "mainly successful" by the ATLAS collaboration. There were difficulties during the execution that were be due to failures in Grid middleware. PanDA, a system for handling distributed production and analysis, was developed in response to the DC2 experience, and it is been used currently in the U.S. for production of data from Cathode Strip Chambers (CSC).

The Core Framework was delivered as a working product (Athena), and the group has shifted largely to performing maintenance and minor upgrades. Input is being sought from the community on the ease of use of Athena, for which management has convened a Usability Task Force.

The Data Management group is concentrating on issues related to the use of ROOT with ATLAS data, including important details of tagged databases, “schema” evolution, and I/O efficiency.

4.2.2 OBSERVATIONS

U.S. ATLAS Core Software has made significant progress since the time of the previous major review in 2005. In particular, a large-scale exercise, Data Challenge 2 (DC2), was carried out and deemed "mainly successful". The committee observes that there were difficulties during this exercise, and that they appear to have been due to failures in Grid middleware, and not due to any particular problems in ATLAS Core Software.

The committee notes that the scheme of "pilot jobs" that PanDA uses for interacting with an underlying batch system seems to complicate the accounting of resources consumed by different users. And because the batch system always appears to be "busy", if only with pilot jobs, this also seems to complicate accounting of the use of facility resources.

4.2.3 RECOMMENDATIONS

The committee made the following recommendations:

- U.S. ATLAS should articulate a plan for verifying that the proposed solution for schema evolution works sufficiently well to meet the needs of the collaboration.

- U.S. ATLAS should periodically revisit the matter of staffing of the analysis-support group to determine whether it is working effectively, especially with regard to the use of software professionals. (What is “schema evolution”?)

5 APPENDIX A

U.S. Department of Energy
and the
National Science Foundation

U.S. LHC Joint Oversight Group

To: Deb Agarwal (LBL), Amber Boehnlein (Fermilab), David Brown (LBL), Richard Dubois (SLAC), Tobias Haas (DESY), Graham Heyes (TJNAF), Albert Lazzarini (Caltech), Mauro Morandin (INFN-Padova), David Morrison (BNL), Doug Olson (LBL), Jon Urheim (Indiana)

Subject: Charge for the February 2006 Review of the U.S. LHC Software and Computing programs

The Joint Oversight Group (JOG) for the U.S. Large Hadron Collider (LHC) Program, supported by the Department of Energy and the National Science Foundation (DOE/NSF), greatly appreciates your willingness to participate in the review of the Software and Computing (S&C) progress and plans for the U.S. LHC Program. The review will take place at the University of California at San Diego on February 7-9, 2006.

The purpose is to evaluate the progress and plans of U.S. ATLAS and U.S. CMS S&C activities in order to assess the effectiveness of the management structures, and to learn whether the S&C activities are sufficiently strong and focused to facilitate the research of U.S. collaborators at the LHC. The review will concentrate on the scope, cost, and schedule of the S&C plans for the period leading up to the turn-on of the LHC, and scrutinize the needs of U.S. ATLAS and U.S. CMS for the initial period of LHC running. To this end, the collaborations will provide plans, including schedules, budgets, risk analysis and contingency plans for the development, deployment, and operation of the U.S. LHC software and computing infrastructure. The information should cover the period FY 2006 through FY 2010, with special emphasis on FY 2007 and 2008, and should be developed for the following two funding scenarios:

- Guidance Level (formerly referred to as "Barebones") as described in the advance material sent to you, with internal allocations for S&C determined by the U.S. collaborations;
- Reduced Level, reduced by 10% from the Guidance Level. For this scenario, the collaborations should present the priorities of the U.S. S&C effort and the impact of reduced funding on both domestic and international schedules and deliverables.

The reviewers should also assess the progress made in implementing recommendations of previous reviews (March 2005 and August 2005). As a guide, we point to the following issues:

Management, External Interactions, and Physics

- Is the overall scope of the U.S. S&C effort well-matched to the needs of the community in exploiting LHC science opportunities?

- Are the estimated costs valid and well-justified in the various cost categories?
- Do the U.S. projects have strong connections to and communications with the international S&C efforts, the Worldwide LHC Computing Grid (WLCG) and the Open Science Grid (OSG)?
- Is risk evaluated and managed adequately? How might the risk of shortfalls in international S&C efforts or projects external to LHC Computing affect U.S. milestones? Are there areas missed in the evaluation of risk?
- Is there adequate contingency to react and adapt to budget and schedule uncertainties? Are the risk-management mechanisms appropriate?
- Is there sufficient communication between U.S.-ATLAS and U.S.-CMS? Are there further areas where common projects could be used to leverage overall resources?
- Assess the effectiveness of the physics-analysis models and whether they take into account the U.S. community's needs within the context of the international collaborations.
- Does management have adequate S&C plans to accommodate new collaborators? Have they developed a reasonable model for the corresponding incremental costs?

Facilities, Grids, and Infrastructure

- Assess the function and scope of the national U.S. LHC computing facilities (Tier-1 centers), their relationship to CERN (Tier-0 center) and to the regional facilities (Tier-2 centers), and whether present plans (hardware, grid software, and networking) are adequate for satisfying the needs as outlined in the experiments' documentation of computing models.
- Do the results of the latest round of data and service challenges lend support to the computing models proposed by U.S. scientists? Are U.S. scientists providing sufficient feedback on problems specific to U.S. involvement?
- Have infrastructure and operating costs of the Tier-1 and Tier-2 facilities been fully considered in their plans? Are there any assumptions that would be regarded as bearing high-risk? Are the estimated costs valid and well-justified?

Core Software

- Is the U.S. carrying its fair burden of the effort in core software, including leadership responsibilities?
- Are the U.S. collaborations sufficiently vigilant in controlling "scope creep"? Is there a well-defined strategy for defining the scope of U.S. participation and for the transition from development to production software?
- Is the U.S. core software portfolio balanced so as to give U.S. researchers a realistic chance for effective participation in the science of the LHC?
- How does the progress in core software measure up to the milestones shown at the previous comprehensive DOE/NSF review of the U.S. program in March 2005? Are U.S. milestones on track and realistic? Is there any critical dependence on international milestones that brings substantial risk to U.S. deliverables?

The review will be chaired by the U.S. LHC Research Associate Program Manager, Saul Gonzalez, with additional program staff from the DOE and the NSF in attendance. You will receive all available documentation at least one week prior to the start of the review. We will appreciate close-out statements following the reviews of both ATLAS and CMS, and more formal written reports within four weeks of the completion of your evaluation. This will provide valuable

and timely input to the agencies and to the experiments. Your reports will also be made available to other DOE and NSF committees that review U.S. ATLAS and U.S. CMS projects.

Again, we wish to express our great appreciation for your willingness to participate in this important activity.

Sincerely,

John R. O'Fallon Co-Chair U.S. LHC Joint Oversight Group Department of Energy	John Lightbody, Jr. Co-Chair U.S. LHC Joint Oversight Group National Science Foundation

cc: Tom Ferbel, SC-25
Aesook Byon-Wagner, SC-25
Moishe Pripstein, NSF/MPS
Miriam Heller, NSF/OCI
Glen Crawford, SC-25
Saul Gonzalez, SC-25
Craig Tull, SC-25
Dan Green, Fermilab
Bob Cousins, UCLA
Jim Shank, Boston University
Lothar Bauerdick, Fermilab
Michael Tuts, Columbia University
Howard Gordon, BNL
Hugh Montgomery, Fermilab
Sam Aronson, BNL

5 APPENDIX B

Review Panel Membership:

D. Agarwal (LBNL)	R. Dubouis (SLAC)	M. Morandin (INFN-Padova)
A. Boehnlein (FNAL)	T. Haas (DESY)	D. Morrison (BNL)
D. Brown (LBNL)	A. Lazzarini (Caltech)	D. Olson (LBL)

Agency Participation:

Glen Crawford (DOE/HEP)
Tom Ferbel (DOE/HEP)
Saul Gonzalez (DOE/HEP, Chair)
Miriam Heller (NSF/OCI)
Moishe Pripstein (NSF/EPP)
Craig Tull (DOE/HEP)

5 APPENDIX C

AGENDA

U.S. LHC Software and Computing Review
February 7-9, 2006
UC San Diego

February 7

CMS

8:30 Executive Session (30', Room TBA)

Plenary Session (1:15, Room TBA)

9:00 U.S. CMS Overview and CMS Computing – L. Bauerdick (30')
9:35 CMS Software and Preparations for Physics – A. Yagil (20')

10:00 Coffee (15')

Plenary Session (1:15, Room TBA)

10:15 U.S. CMS S&C Status and Plans – I. Fisk (35')
10:55 U.S. CMS Tier-2 Program – K. Bloom (20')
11:20 DISUN Status and Plans – F. Würthwein (20')

11:45 Executive Session (15')

12:00 Lunch

13:00 Parallel Session (3:00)

Session:	Core Software	Management, External Interactions, Physics	Facilities, Grids, Infrastructure
Room:	TBA	TBA	TBA
CMS Participants:	Sexton-Kennedy, Yagil, Clare	Bauerdick, Fisk, Green, Yagil, Cousins, Tschirhart, Montgomery	Bakken, Bloom, Würthwein, Pordes, Fisk

16:00 Coffee (15')
16:15 Executive Session and Questions for CMS (2:15, Room TBA)
18:30 Adjourn

February 8

CMS

8:30 CMS Response to questions from committee (30', Room TBA)
9:00 Executive session (2:30, Room TBA)
11:30 CMS Closeout (30', Room TBA)
12:00 Lunch

February 8

ATLAS

13:00 Executive Session (15', Room TBA)

Plenary Session (2:30, Room TBA)
13:15 ATLAS Introduction – J. Shank (5')
13:20 ATLAS Overview – D. Quarrie (20')
13:45 Program Status – J. Shank (35')
14:25 Software Status – S. Rajagopalan (35')
13:05 Facilities – B. Gibbard (35')

15:45 Coffee (15')

16:00 Executive Session (15')
16:15 Parallel Session (1:30)

Session:	Core Software	Management, External Interactions, Physics	Facilities, Grids, Infrastructure
Room:	TBA	TBA	TBA
ATLAS Participants	16:15 DDM – Wenaus (30') 16:45 PanDa – De (30')	16:15 Management --Shank, Gordon, Tuts, Huth (1:30)	16:15 Tier-1 – Gibbard (20') 16:45 Tier-2 – Gardner (20') 17:15 Networking – McKee (20')

17:45 Executive Session and Questions for ATLAS (45', Room TBA)
18:30 Adjourn

February 9

ATLAS

8:30 Parallel Session (1:30)

Session:	Core Software	Management, External Interactions, Physics	Facilities, Grids, Infrastructure
Room:	TBA	TBA	TBA
ATLAS Participants	8:30 Core SW – Calafiura (15') 8:50 DB – Malon (15') 9:20 Analysis Support – Willocq (30')	8:30 Management -- Shank, Gordon, Tuts, Huth (1:30)	8:30 OSG/WLCG – Huth (30')

10:00 Executive Session and Questions for ATLAS (1:45)

11:45 Lunch

14:00 ATLAS answers to committee's questions (30')
14:30 Executive Session (2:30)
17:00 ATLAS Closeout (30')
17:30 Adjourn